

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
КАФЕДРА ТЕОРИИ СИСТЕМ УПРАВЛЕНИЯ ЭЛЕКТРОФИЗИЧЕСКОЙ АППАРАТУРОЙ

**Тонеева Дарья Владимировна**

**Выпускная квалификационная работа бакалавра**

**Экспертная система для диагностики заболеваний,  
построенная на базе совокупности  
дифференциально-диагностических признаков**

Направление 010900  
«Прикладные математика и физика»

Научный руководитель,  
кандидат физ.-мат. наук,  
доцент  
Головкина А.Г.

Санкт-Петербург  
2017

# Содержание

|  |    |
|--|----|
| Введение.....  | 3  |
| Постановка задачи.....   | 5  |
| Обзор экспертных систем.....   | 7  |
| Глава 1. Построение медицинской экспертной системы.....  | 15 |
| 1.1. Алгоритм формирования входных параметров .....  | 15 |
| 1.2. Модель принятия решений на примере заболевания артрита.....                               | 19 |
| Глава 2. Программная реализация экспертной системы для ранней<br>диагностики заболеваний ..... | 22 |
| 2.1. База данных симптомов.....  | 22 |
| 2.1. Структура программы.....  | 23 |
| 2.3. Пользовательский интерфейс .....  | 25 |
| Выводы .....   | 28 |
| Заключение .....   | 29 |
| Список литературы .....  | 30 |

## Введение

Экспертная система (ЭС) – это компьютерная система, содержащая в себе знания специалистов в конкретной предметной области, и которая на основе этих знаний способна к самостоятельному принятию решений. ЭС получила широкое распространение в области медицины в качестве помощника врача при постановке предварительного диагноза.

Специалист может обратиться за помощью к ЭС в том случае, когда задача четко не формализована, и алгоритм ее однозначного решения труднореализуем. Также актуальным является применение медицинских ЭС при ранней диагностике заболеваний, когда анамнез пациента может содержать перечень сильно коррелирующих или взаимоисключающих симптомов. Тогда появляется необходимость в разработке механизма четкого разграничения симптомов и заболеваний. Здесь значительную помощь может оказать система поддержки и принятия решений (СППР), рассматриваемая как одно из направлений медицинских ЭС.

Выпускная квалификационная работа посвящена разработке и реализации ЭС для ранней диагностики заболеваний. Данная ЭС построена на базе совокупности дифференциально-диагностических признаков, имеет математический аппарат механизма принятия решений и адаптируемый пользовательский интерфейс. Последнее означает, что ЭС может быть использована для широкого круга патологий. Работа ЭС показана на примере диагностики артрита. Программная разработка была выполнена на языке программирования C#.

Выбор артрита в качестве заболевания для демонстрации работы программы связан не только с необходимостью получения требуемых данных из достоверных источников. Также это обусловлено тем, что артрит является довольно распространенным заболеванием, а его диагностика –

непростой задачей.

Артрит – это заболевание суставов воспалительного характера. Наиболее распространенным из всех его типов является ревматоидный артрит (РА). По прогнозам Всемирной организации здравоохранения к 2020 г. РА будет обуславливать высокую частоту инвалидизации пациентов (20-60%), что может привести к ранней потере трудоспособности населения [1]. Подагрический артрит (ПА) также является одним из видов ревматоидных заболеваний наряду с РА и имеет схожие симптомы на хронической стадии. Третий тип ревматоидных заболеваний, который по симптомам на ранней стадии имеет сходства с РА и с ПА, это реактивный артрит (РЕА). Несмотря на то, что причиной РЕА является внесуставная инфекция, до ее выявления болезнь протекает по аналогии с другими видами артрита.

В связи с этим возникает сложность ранней диагностики заболеваний и необходимость в построении диагностической экспертной системы.

В выпускной квалификационной работе рассматриваются перечисленные разновидности артрита, как наиболее часто встречающиеся у пациентов, и которые имеют схожие симптомы на ранней стадии развития. Информация по частоте встречаемости характерных симптомов были взяты из медицинских источников, также были учтены рекомендации по проведению ранней диагностики артрита, предоставленные поликлиникой №28 города Санкт-Петербурга.

## Постановка задачи

Экспертные системы играют очень важную роль в развитии здравоохранения в целом, и служат для того, чтобы обеспечивать систематическую помощь медицинским службам в случае возникновения спорных и проблемных ситуаций в вопросах диагностирования и лечения пациентов. Медицинские ЭС способны решать не только задачи диагностики, но и помогать в прогнозировании течения болезней, в выборе тактики их лечения [2]. Значимость использования медицинских ЭС возрастает в тех случаях, если:

- симптоматическая база данных очень велика для каждого из прогнозируемых заболеваний;
- исходные данные (симптомы) сильно коррелируют между схожими заболеваниями;
- специалист сталкивается с трудностью в построении алгоритма однозначного решения данной задачи (постановки предварительного диагноза).

**Цель данной работы** – разработка и реализация экспертной системы для ранней диагностики различных заболеваний, включающей в себя:

- механизм обращения к базе данных симптомов по каждому из коррелирующих заболеваний;
- алгоритм формирования входных параметров ЭС;
- метод принятия решений на основе математического аппарата вывода.

Для успешного использования данной ЭС она должна обладать следующими свойствами:

- иметь возможность подключения базы данных симптомов различных заболеваний;

- включать в себя адаптивный пользовательский интерфейс, позволяющий применять разработанное программное средство для диагностики широкого круга патологий;
- производить раннюю диагностику заболевания по совокупности дифференциально-диагностических признаков;
- обеспечивать специалиста достоверными результирующими данными о наличии диагностируемого заболевания в процентном соотношении;
- программа должна быть верифицирована на примере диагностирования разновидностей артрита.

## Обзор экспертных систем

Разработка экспертных систем является одним из направлений исследований искусственного интеллекта. Данные системы позволяют решать труднореализуемые для эксперта-человека задачи и одновременно получать качественные результаты, которые соизмеримы с результатами самого эксперта. ЭС способны давать советы, проводить консультации, анализировать данные и принимать самостоятельные решения. Экспертные системы широко применяются в таких областях, как торговля, финансы, нефтяная и газовая промышленность, космос, образование, медицина, энергетика, химия, транспортная логистика и другие [3].

В работе [4] описана работа экспертной системы для анализа различных организаций. Одним из преимуществ данной ЭС является то, что помимо количественных показателей, таких как рентабельность, общая прибыль, она способна также учитывать качественные коэффициенты: положение на рынке, организация персонала, деловой репутации, гибкости маркетинга и другие. Примером системы из сферы финансового анализа служит ЭС FINEVA (FINancial EVAuation – финансовая оценка), разработанная в 1997 г., которая выявляет сильные и слабые стороны компаний посредством качественной оценки и интерпретации финансовых коэффициентов.

В настоящее время деятельность любого бизнеса непосредственно связана с информационными технологиями. Важно непрерывно следить за множеством бизнес-операций: заказ товара, отгрузка, предоставление услуг. Информация стала одним из самых ценных ресурсов организаций. Таким образом, ЭС стали активно использоваться в качестве инструмента информационной безопасности и защиты [5]. Одним из примеров таких ЭС служит экспертная система аудита безопасности (AudES), предназначенная для автоматизации некоторых аудиторских процедур [6].

Интеллектуальные экспертные системы получили самое широкое применение и в области медицины в качестве помощника при планировании лечения и прогнозировании рисков развития осложнений. Это связано с тем, что в экстренных случаях жизнь пациента зависит от сроков постановки диагноза и выбора схемы лечения. Однако врачу для этого может потребоваться время, иногда достаточно продолжительное, особенно в случае спорного диагноза или присутствии в анамнезе пациента хронических заболеваний, способных вызвать развитие осложнений в послеоперационном периоде. В то же время качественная медицинская ЭС способна предоставить адекватные результаты, по точности сопоставимые с назначениями эксперта, причем в самые короткие сроки.

Также медицинские ЭС могут оказаться полезными не только для экстренной, но и для ранней диагностики заболеваний. Подобные системы позволяют выявить начало заболевания, давая возможность своевременно начать терапию, тем самым сократить сроки предполагаемого лечения и даже, возможно, предотвратить развитие болезни. В связи с вышеизложенным роль ЭС в медицине так важна, а проблема создания качественной медицинской ЭС актуальна [3].

Первая медицинская ЭС MYCIN была разработана в начале 1970-х годов. Система служит для диагностики и лечения инфекционных заболеваний. *Пользовательский интерфейс MYCIN* представляет собой «опросник» для пациента, блок *механизма принятия решений* – прямой вывод с использованием правил нечеткой логики и обратный вывод. БДС, которой располагает MYCIN, постоянно расширяется, благодаря этому ЭС применяется также в других областях медицины и служит прототипом для многих медицинских систем. Одной из них является ЭС PUFF, которая применяется в исследовании легочных заболеваний.

В настоящее время существует ряд IT-компаний, специализирующихся на разработке медицинских экспертных систем. Одной из них является Socmedica, реализовавшая технологию моделирования медицинских знаний,



на основе которой Научные Центры РФ формируют Объединенную Базу Медицинских Знаний – УМКВ [7]. Socmedica ведет разработки экспертных систем по прогнозированию рисков развития заболеваний, осложнений и эффективности лечения, по ранней диагностике, по мониторингу состояния здоровья пациента и другие.

Для более эффективной работы медицинских экспертных систем компания «АБ Систем» создала специализированный программный комплекс «ЭПИКО – Учетная медицинская система», который содержит данные историй болезней пациентов. Благодаря учетной системе работа ЭС может быть автоматизирована, что позволяет снизить риск возникновения ошибок и ускорить процесс принятия решений [8].

Структура медицинской ЭС включает в себя пользовательский интерфейс, базу данных симптомов (БДС) и механизм принятия решений, позволяющий системе принимать решение о постановке предварительного диагноза без помощи специалиста. Такой механизм может быть *логическим*, т.е. представлять собой совокупность формализованных фактов и правил логического вывода (базу знаний), *математическим*, т.е. включающим себя математический аппарат вывода результирующих данных или *гибридным*, т.е. объединением первых двух методов. На рис.1 представлена структура ЭС.

При создании БДС специалисты в области медицины могут столкнуться со следующей проблемой: одному симптому могут соответствовать сразу несколько наименований в зависимости от источников данных. Такая ситуация может возникнуть, когда над составлением БДС работает не один эксперт и их терминология различна. В случае, когда симптомов немного, эта проблема не столь заметна. Однако когда объем БДС достаточно велик, база данных симптомов занимает очень большой объем памяти и обращение ЭС к ней может занимать значительное количество времени, что недопустимо.

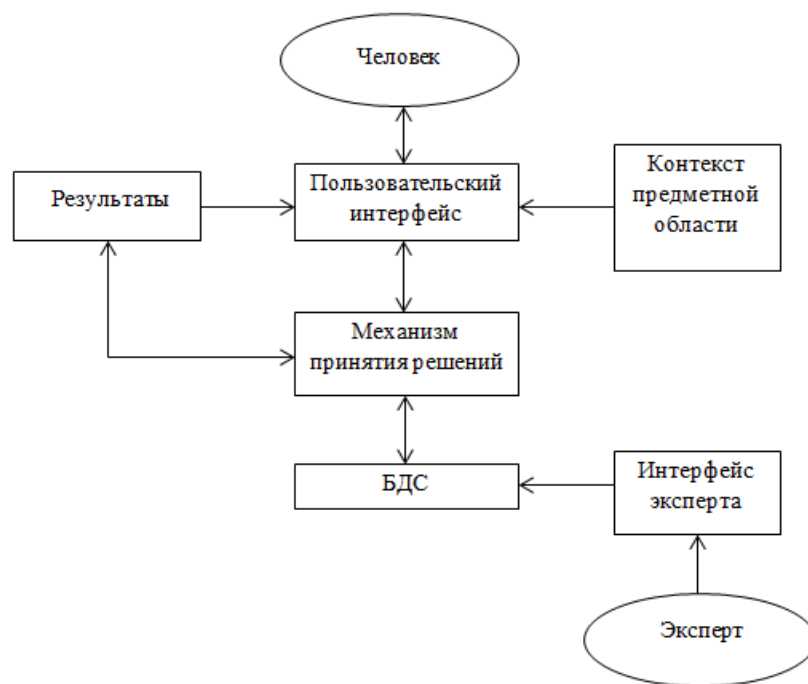


Рис. 1 Схема экспертной системы

В работе [9] предложено формализовать клиническую информацию для оптимизации диагностического процесса так, чтобы за каждым симптомом находился свой уникальный код. То есть множеству текстовых шаблонов будет сопоставлено одно явление – симптом, соответствующий своему уникальному коду. Этого можно достичь, если сделать акцент не на различных наименованиях одного симптома, а на терминах, описывающих его. Например, симптомы «послабляющая температура» и «ремитирующая температура» описывают один симптом и должны иметь единый уникальный код. Такое отношение между именами, когда наименования ссылаются на один и тот же объект, называется *кореферентным* или *референциальным тождеством*.

Всемирная организация здравоохранения (ВОЗ) сделала серьезный шаг в решении проблемы с терминологией заболеваний, разработав международную классификацию болезней (МКБ). Данная классификация раз в десять лет подлежит пересмотру и дополнению новых заболеваний. Для того чтобы аналогичным образом решить проблему с формализацией базы

данных симптомов, ВОЗ создает международную классификацию функционирования, ограничений жизнедеятельности и здоровья (МКФ).

МКФ, как и другие международные классификации, обеспечивает общие правила кодирования широкого круга информации, связанной со здоровьем (например, функционирование и ограничение жизнедеятельности и др.) и использует стандартизированный общий язык, позволяющий общаться по проблемам, связанными со здоровьем во всем мире в различных дисциплинах и отраслях науки. МКФ является классификацией доменов здоровья и доменов, связанных со здоровьем. В работе [10] разработана база данных симптомов на основе данных МКФ для ведения первичного приема детей раннего возраста с различными нарушениями.

Для хранения большого объема данных чаще всего используются реляционные базы данных. В качестве системы управления базами данных (СУБД) в работе [10] применяется СУБД Microsoft Access, которая относится к реляционным системам. Связь базы данных с программой Delphi 7 была реализована с помощью механизма ActiveX Data Object (ADO). Модель базы данных состоит из 25 основных таблиц. Для манипулирования данными используются запросы, реализованные на языке SQL в среде Delphi.

Такой подход позволил авторам также осуществлять выборку информации из базы данных и совершать другого рода манипуляции с хранящимися данными в программной среде, созданной на языке Delphi 7.

Следующий этап разработки медицинской ЭС является реализация механизма принятия решений. В работе [11] подробно описан метод использования гибридного механизма. В этом случае для постановки диагноза применяются деревья принятия решений (ДПР). ДПР – это направленный граф, у которого конечные вершины (вершины графа, у которых нет дуг) представляют собой вывод (некоторые альтернативы выбора). Оставшиеся промежуточные вершины графа содержат в себе условие перехода к другим вершинам (см. рис. 2 и рис. 3). Для того чтобы использовать ДПР наиболее эффективно, авторы выбрали в качестве

алгоритма принятия решений ЭС алгоритм генетического программирования (ГП). Для успешного использования ГП необходимо выполнить два шага: представить объекты поиска, как иерархические деревья, и формализовать процедуру их оценки.

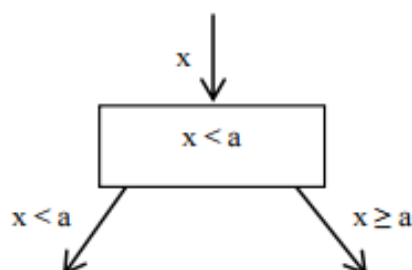


Рис. 2 Функциональный узел ДПР:  $x$  – входящая переменная;  $a$  – параметр узла

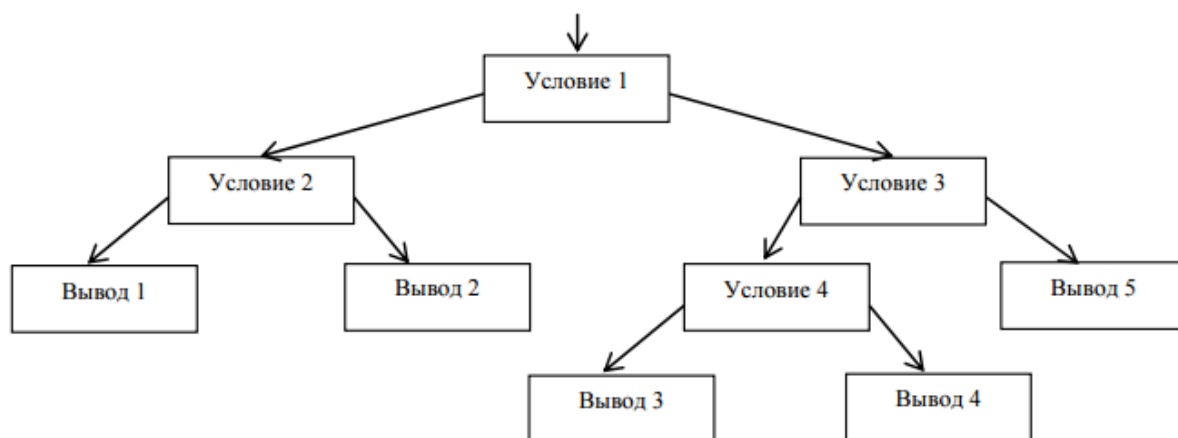


Рис. 3 Дерево принятия решений

Алгоритм называется генетическим, так как его основной задачей является «скрещивание», комбинирование пар вариантов решения. Далее происходит селекция – отбор, и формирование нового поколения с помощью мутации, которое наследует некоторые признаки предыдущего. Этап мутации есть не что иное, как случайные изменения в ДПР с достаточно малой вероятностью. Функциональная вершина (промежуточная вершина графа) заменяется на случайно выбранную функциональную вершину, а терминальная вершина (конечная вершина графа) – на терминальную [11].

Далее снова происходит скрещивание и отбор. Так продолжается до тех пор пока:

А) число циклов (поколений) не достигнет заранее выбранного максимума;

Б) не исчерпано время работы программы.

Результатом работы ЭС на основе генетического программирования будет формирование дерева принятия решений с наибольшей пригодностью.

Экспертная система, реализованная на основе генетического программирования, имеет свои недостатки:

- 1) на «селекцию» – перебор всех альтернативных вариантов программе требуется много времени;
- 2) число поколений (циклов программы) ограничено – в результате решение задачи будет приближенным.

Применение *математического аппарата* механизма принятия решений позволяет избежать этих недостатков. В работе [12] приведен наглядный пример эффективного использования математического метода принятия решения для БДС, построенной на базе совокупности дифференциально-диагностических признаков. Однако данный метод был разработан лишь для диагностики видов рака щитовидной железы.

Медицинская ЭС в этом случае сопоставляет матрицу симптомов пациента  $X$  эталонным матрицам заболеваний  $S$ . В самих матрицах хранится массив *весовых коэффициентов*  $k$  по каждой *частоте встречаемости*  $f$  симптома в конкретном заболевании. Частота встречаемости  $f$  определяется эмпирически по выборке пациентов с помощью подсчета встречаемости каждого из симптомов:

$$f = \frac{N_+}{N_+ + N_-} \times 100\% ,$$

где  $N_+$  – пациенты из выборки с наличием данного симптома,  $N_-$  – пациенты, у которых данный симптом отсутствует.

При использовании математического подхода принятия решений, весовой коэффициент  $k_i$   $i$ -го симптома рассчитывается следующим образом:

$$k_i = \frac{f_i}{\sum f_j} \cdot 100\%,$$

где  $f_i$  – частота встречаемости  $i$ -го симптома, соответствующая данному весовому коэффициенту [12].

Далее для постановки предварительного диагноза рассчитывается *диагностический индекс*  $D$  по формуле:

$$D = \sum k_i p_i, \quad (1)$$

где

$$p_i = \begin{cases} 0, & \text{значения } i - \text{х ячеек матрицы } X \text{ и матрицы } S \text{ не равны;} \\ 1, & \text{значения } i - \text{х ячеек } X \text{ и } S \text{ совпадают.} \end{cases}$$

Диагностический индекс позволяет рассчитать вероятность наличия того или иного заболевания в процентном отношении.

В выпускной квалификационной работе для разработки ЭС был использован *математический метод* принятия решений. Благодаря *адаптируемому пользовательскому интерфейсу*, данная ЭС позволяет получать результаты для широкого круга патологий и вычислять вероятность того или иного заболевания в процентном отношении.

# Глава 1. Построение медицинской экспертной системы

## 1.1. Алгоритм формирования входных параметров

В работе [13] подробно описан алгоритм построения экспертной системы на базе совокупности дифференциально-диагностических признаков. Его суть состоит в следующем.

Пусть имеется  $n$  больных в контрольной группе, у которых установлено  $m$  диагнозов. Среди  $n$  больных  $p_i$  имеют диагноз  $i$ , то есть:

$$\begin{cases} p_1 - \text{имеют диагноз } 1; \\ p_2 - \text{имеют диагноз } 2; \\ \dots \\ p_m - \text{имеют диагноз } m, \end{cases}$$

причем некоторые больные могут с некоторой вероятностью иметь как диагноз  $p_i$ , так и диагноз  $p_j$ , и таким образом, имеет место неравенство:

$$p_1 + \dots + p_i + \dots + p_m \geq n.$$

Рассмотрим диагноз  $D$ . По результатам обследования у  $p_D$  больных имеются симптомы:  $l_1 \dots l_s$  – набор дифференциально-диагностических признаков (симптомов) болезни  $D$ , т.е. образуется вектор симптомов:

$$p_D(l_1, l_2 \dots l_s).$$

Симптомы, в свою очередь, принадлежат подгруппе некоторой группы признаков. Например, группа признаков «общеклинические исследования» будет включать в себя следующие подгруппы: клинический анализ крови с лейкоцитарной формулой и СОЭ, копрограмму, анализ кала на цисты и вегетативные формы простейших, общий анализ мокроты и другие. Группа «биохимические анализы» будет включать в себя подгруппы: биохимический анализ крови, билирубин, глюкозу в плазме, белок общий в сыворотке и другие.

Рассчитаем частоту встречаемости  $f_{Di}$ , к примеру, признака  $l_i$ ,  $i = \overline{1, s}$ , которая определяется эмпирически путем взятия выборки пациентов и подсчета встречаемости каждого из симптомов [14]. Стоит отметить, что чем больше будет объем выборки, тем точнее будет показатель частоты встречаемости. Затем для контрольной группы  $p_D$  больных с диагнозом  $D$  рассчитаем *весовой коэффициент*  $k_{Di}$  по формуле:

$$k_{Di} = \frac{f_{Di}}{\sum f_{Dj}} \cdot 100\%, \quad (2)$$

где  $f_{Di}$  – частота встречаемости  $i$ -го симптома диагноза  $D$ , соответствующая данному весовому коэффициенту.

Далее построим *эталонный вектор патологии* для болезни  $D$  *логического типа*. Ячейка вектора имеет значение *true*, если данное заболевание имеет данный симптом, и значение *false* иначе [15].

В результате можно составить *Таблицу 1*, в которую входят *эталонный вектор* болезни  $D$ , частота встречаемости и весовой коэффициент каждого симптома [10].

Таблица.1.

Эталонные ответы да/нет, частота встречаемости и весовые коэффициенты дифференциально-диагностических признаков

| Признаки |                   | Патология                 |  |  |
|----------|-------------------|---------------------------|--|--|
|          |                   | Болезнь D                 |  |  |
| Группа   | Подгруппа         | Эталонные<br>ответы (+/-) | Частота<br>встречаемости<br>$f_{Di}, \%$ | Весовой<br>коэффициент<br>$k_{Di}, \%$ |
| Группа 1 | Признак $l_1$     | +                         | $f_{D1}$                                 | $k_{D1}$                               |
|          | Признак $l_2$     | -                         | 0  | 0                                      |
|          | Признак $l_3$     | +                         | $f_{D3}$                                 | $k_{D3}$                               |
| Группа 2 | Признак $l_4$     | +                         | $f_{D4}$                                 | $k_{D4}$                               |
|          | Признак $l_5$     | -                         | 0  | 0                                      |
|          | Признак $l_6$     | +                         | $f_{D6}$                                 | $k_{D6}$                               |
| ...      | ...               | ...                       | ...                                      | ...                                    |
| Группа k | Признак $l_{k+t}$ | -                         | 0  | 0                                      |



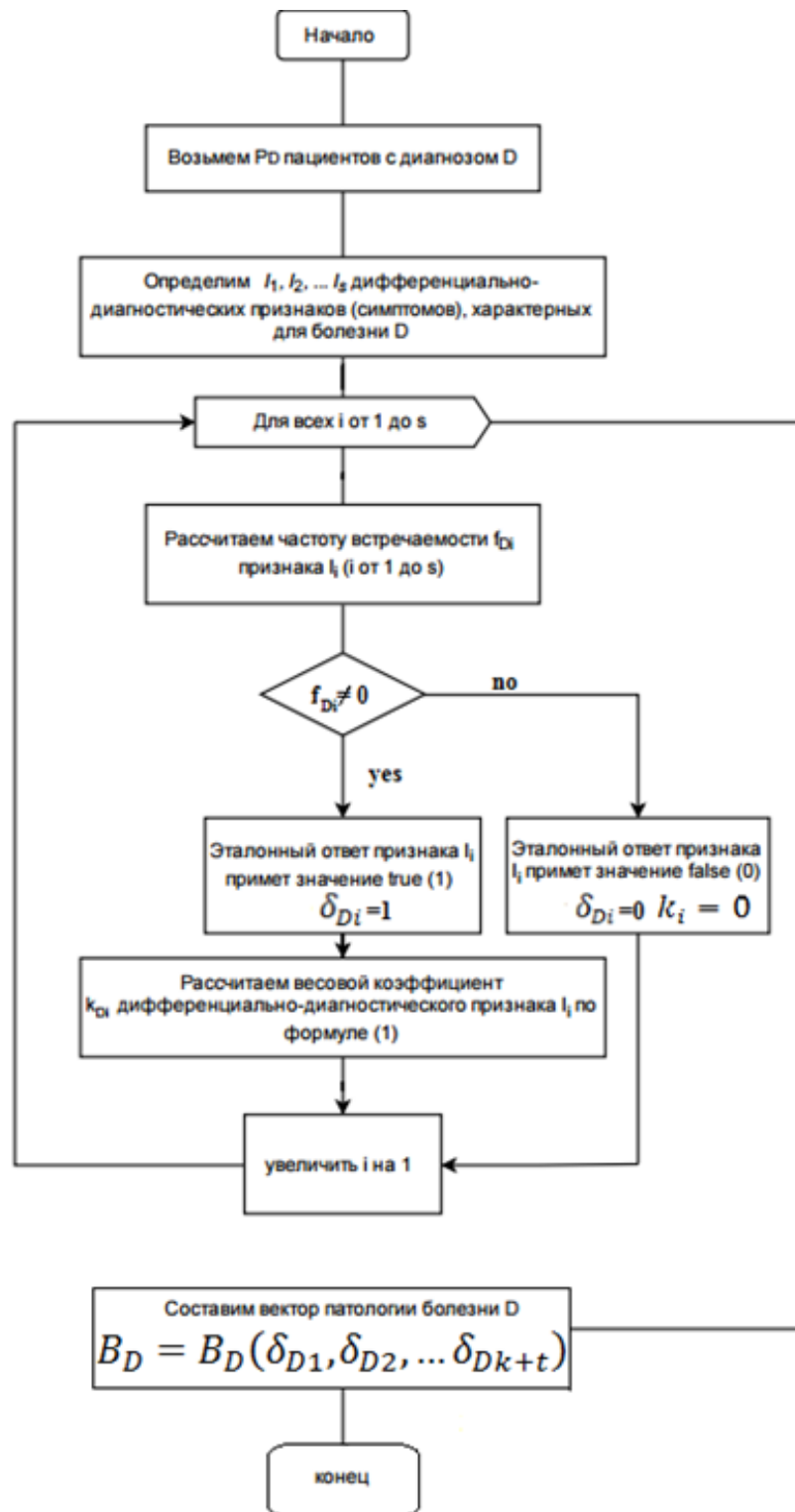


Рис.4 Алгоритм построения эталонного вектора патологии болезни

Таким образом, любому диагнозу  $D$  однозначно соответствует определенный набор симптомов – признаков, которые можно представить в виде вектора соответствия  $B_D$ , где «1» и «0» ставится в соответствие «+» и «-» (см. рис.4):

$$B_D = B_D(\delta_{D1}, \delta_{D2}, \dots \delta_{Dk+t}),$$

где  $\delta_{ij} = \begin{cases} 1, & \text{если данное заболевание имеет данный симптом } l_i; \\ 0, & \text{иначе.} \end{cases}$

Отвечая на вопросы о наличии или отсутствии симптома у больного, врач формирует *вектор соответствия больного j* по аналогии с работой [16]:

$$j = j(\delta_{j1}, \delta_{j2}, \dots \delta_{jk+t}),$$

где  $\delta_{ij} = \begin{cases} 1, & \text{если пациент имеет данный симптом;} \\ 0, & \text{иначе.} \end{cases}$

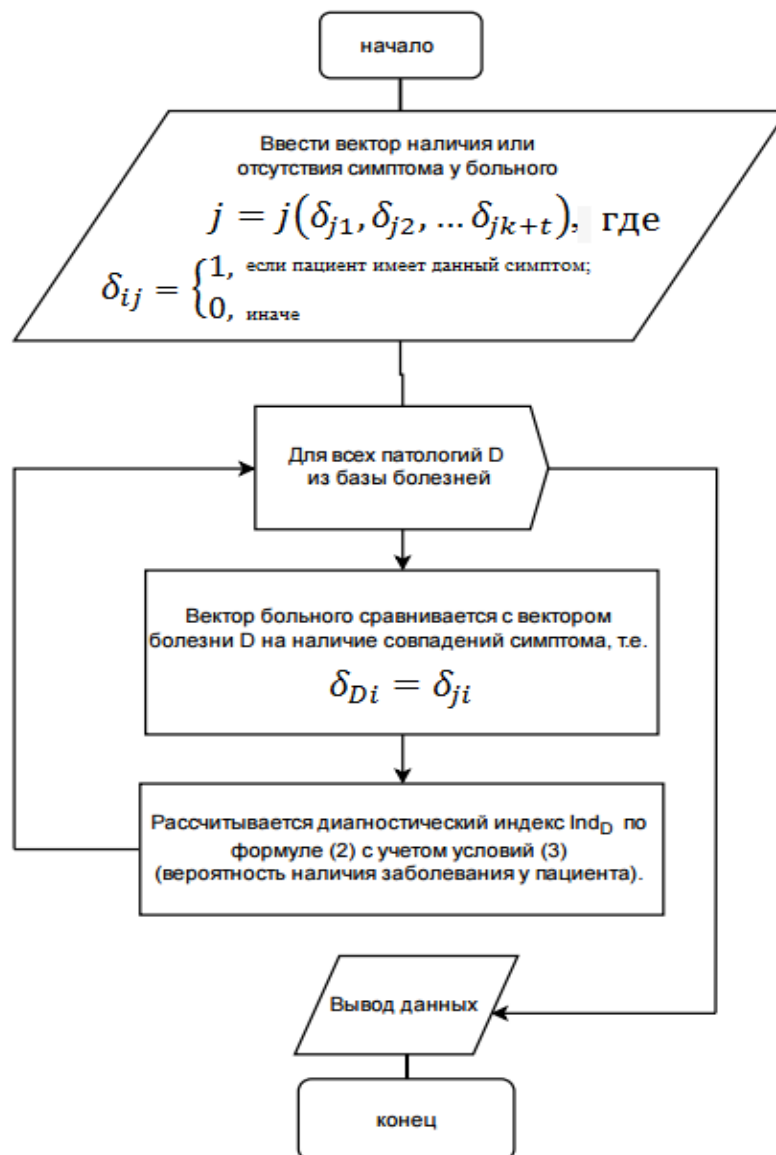


Рис. 5 Алгоритм принятия решений диагностики заболевания

Введем операцию  $\&\&$  такую что:

Для  $\forall$  векторов  $\vec{a} = (a_1, a_2 \dots a_n)$  и  $\vec{b} = (b_1, b_2 \dots b_n)$  одинаковой размерности:

$$\vec{a} \&\& \vec{b} = \vec{c},$$

где  $\vec{c} = (a_1 \& b_1, a_2 \& b_2 \dots a_n \& b_n)$ ,  $\&$  – логическая операция И.

Тогда

$$Ind_D = (\vec{c}, \vec{k}),$$

где  $\vec{c} = j \&\& B_D$ ,  $\vec{k} = (k_1, k_2 \dots k_n)$  – вектор весовых коэффициентов.

Таким образом, для каждого пациента рассчитывается вероятность наличия того или иного заболевания в процентном отношении (см. рис.4).

## 1.2. Модель принятия решений на примере заболевания артрита

Для построения экспертной системы на основе описанного выше алгоритма были использованы материалы из работы [17]. Отбирались симптомы и их частота встречаемости из выборки пациентов на основе работ по диагностике артрита [18], [19] и [20]. В работе [18] разработаны критерии, позволяющие выявить ревматоидный артрит на ранней стадии. В качестве признаков были взяты те показатели, которые могли зафиксировать наличие прогрессирующего воспаления и неприятные прогностические факторы.

В работе [11] метод ранней диагностики подагрического артрита полагается на типичные клинические признаки болезни. Если итоговая сумма баллов пациента ниже 4, то вероятность подагрического артрита будет равна только 2%, если сумма входит в диапазон 4-8 баллов, тогда вероятность 31%, больше 8 баллов – 80%.

Реактивный артрит является системным заболеванием и имеет совокупность внесуставных проявлений [12]. Реактивный артрит развивается

после перенесенной больным инфекции, поэтому симптомами могут быть различные поражения глаз, мочеполового тракта, желудочно-кишечного тракта, кожи, сердечно - сосудистой системы.

Данные по симптомам и соответствующей им частоте встречаемости занесены в *Таблицу 2* [14]. Весовой коэффициент  $k_i$  рассчитан по формуле (2):

*Таблица 2.*

Эталонные ответы да/нет, частота встречаемости и весовые коэффициенты дифференциально-диагностических признаков артрита.

| Признаки  |   | Патология                |                                       |                                     |                          |                                       |                                     |                          |                                       |                                     |
|---|---|--------------------------|---------------------------------------|-------------------------------------|--------------------------|---------------------------------------|-------------------------------------|--------------------------|---------------------------------------|-------------------------------------|
|   |   | Ревматоидный артрит      |                                       |                                     | Подагрический артрит     |                                       |                                     | Реактивный артрит        |                                       |                                     |
| Группа  | Подгруппа   | Эталонные<br>ответы(+/-) | Частота<br>встречаемости<br>$f_i, \%$ | Весовой<br>коэффициент<br>$k_i, \%$ | Эталонные<br>ответы(+/-) | Частота<br>встречаемости<br>$f_i, \%$ | Весовой<br>коэффициент<br>$k_i, \%$ | Эталонные<br>ответы(+/-) | Частота<br>встречаемости<br>$f_i, \%$ | Весовой<br>коэффициент<br>$k_i, \%$ |
| Клинические признаки поражения суставов (припухлость, воспаление) | 1 Крупный сустав (покраснение)  | -                        | 0                                     | 0                                   | +                        | 10                                    | 7,69                                | -                        | 0                                     | 0                                   |
|   | 2-10 крупных суставов   | +                        | 10                                    | 5,43                                | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   |
|   | 1-3 мелких сустава (без учета крупных)                                    | +                        | 20                                    | 10,87                               | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   |
|   | 4-10 мелких сустава (без учета крупных)                                   | +                        | 30                                    | 16,3                                | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   |
|   | >10 суставов (как минимум 1 мелкий сустав)                                | +                        | 50                                    | 27,17                               | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   |
|   | Вовлечение первого плюснефалангового сустава                              | -                        | 0                                     | 0                                   | +                        | 25                                    | 19,23                               | -                        | 0                                     | 0                                   |
|   | Энтезопатии( боль в пятке, боль в проекции большеберцового бугра)         | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   | +                        | 30                                    | 8,89                                |
|   | Воспалительная боль в нижнем отделе спины: сакроилеит                     | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   | +                        | 31                                    | 9,33                                |
|   | Воспалительная боль в нижнем отделе спины: спондилит                      | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   | +                        | 19                                    | 5,63                                |
|   | Воспаление связок и сухожилий в месте их прикрепления к седалищному бугру | -                        | 0                                     | 0                                   | -                        | 0                                     | 0                                   | +                        | 22                                    | 6,52                                |
| Атака артрита   | Достижение максимальных проявлений артрита во время атаки за один         | +                        | 1                                     | 0,54                                | +                        | 5                                     | 3,85                                | -                        | 0                                     | 0                                   |

|                                    |   |   |    |       |   |    |       |   |    |       |
|------------------------------------|---|---|----|-------|---|----|-------|---|----|-------|
|                                    | день  |   |    |       |   |    |       |   |    |       |
|                                    | Одна или более атак артрита в анамнезе  | + | 3  | 1,63  | + | 20 | 15,38 | - | 0  | 0     |
|                                    | Периферический артрит (несимметричный, олигоартрит)                               | - | 0  | 0     | - | 0  | 0     | + | 22 | 6,52  |
| Тесты на РФ и АЦЦП                 | Отрицательны  | - | 0  | 0     | - | 0  | 0     | - | 0  | 0     |
|                                    | Слабо позитивны для РФ или АЦЦП (превысили границу нормы, но менее, чем в 3 раза) | + | 20 | 10,87 | - | 0  | 0     | - | 0  | 0     |
|                                    | Высоко позитивны для РФ или АЦЦП (более, чем в 3 раза превысили границу нормы)    | + | 30 | 16,3  | - | 0  | 0     | - | 0  | 0     |
| Показательные<br>Анализ            | Норма по СОЭ (Скорость оседания эритроцитов) и СРБ (С-реактивный белок)           | - | 0  | 0     | - | 0  | 0     | - | 0  | 0     |
|                                    | Повышение СОЭ или уровня СРБ  | + | 10 | 5,43  | - | 0  | 0     | - | 0  | 0     |
|                                    | Мочевая кислота >6,0 мг/дл (360 мкмоль/л)   | - | 0  | 0     | + | 35 | 26,92 | - | 0  | 0     |
| Переносимые процессы и/или болезни | Синовит < 6 нед.  | - | 0  | 0     | - | 0  | 0     | - | 0  | 0     |
|                                    | Синовит ≥ 6 нед.  | + | 10 | 5,43  | - | 0  | 0     | - | 0  | 0     |
|                                    | Гипертензия и/или одна или более сердечно-сосудистых заболеваний                  | - | 0  | 0     | + | 15 | 11,54 | - | 0  | 0     |
|                                    | Конъюнктивит  | - | 0  | 0     | - | 0  | 0     | + | 30 | 8,89  |
|                                    | Уретрит, простатит  | - | 0  | 0     | - | 0  | 0     | + | 80 | 23,7  |
|                                    | Эндоскопические признаки поражения кишечника                                      | - | 0  | 0     | - | 0  | 0     | + | 47 | 14,07 |
|                                    | Бленноррагическая кератодермия  | - | 0  | 0     | - | 0  | 0     | + | 17 | 5,19  |
|                                    | Эрозивный круговидный баланит (4–20%)   | - | 0  | 0     | - | 0  | 0     | + | 12 | 3,56  |
|                                    | Язвы слизистой оболочки полости рта   | - | 0  | 0     | - | 0  | 0     | + | 7  | 2,22  |
|                                    | Гиперкератоз ногтей   | - | 0  | 0     | - | 0  | 0     | + | 9  | 2,67  |
|                                    | Нарушения проводимости по данным ЭКГ  | - | 0  | 0     | - | 0  | 0     | + | 9  | 2,81  |
|                                    | Мужской пол   | - | 0  | 0     | + | 20 | 15,38 | - | 0  | 0     |
|                                    | Женский пол   | - | 0  | 0     | - | 0  | 0     | - | 0  | 0     |
| Итого                              |   |   |    | 100   |   |    | 100   |   |    | 100   |

## Глава 2. Программная реализация экспертной системы для ранней диагностики заболеваний

### 2.1. База данных симптомов

Работа ЭС опирается на базу данных симптомов (БДС). БДС может быть сформирована при опросе экспертов-медиков в данной области, путем статистической обработки данных медицинского учреждения или взята из открытых литературных источников. БДС представляет собой совокупность дифференциально-диагностических признаков, соответствующих частот их встречаемости и весовых коэффициентов для каждого из диагностируемых заболеваний.

В данной работе БДС сформирована для различных видов артрита. Структура этих данных позволяет представить их в табличном виде. Пользователь, а именно специалист (врач или медработник), участвует в составлении БДС по заболеваниям артрита, проверяет данные на достоверность в силу своих медицинских знаний. Данные заполняются в некоторой доступной и удобной специалисту программе по работе с электронными таблицами, из которой далее производится преобразование данных в CSV формат.

CSV файл (Comma Separated Values file) – текстовый формат, предназначенный для представления табличных данных. Структура файла формата csv имеет следующую спецификацию:

- каждая строка файла — это *одна строка* таблицы;
- разделителем значений колонок является символ запятая.

Однако в русской и других локалях (*набор параметров, определяющий региональные настройки пользовательского интерфейса*) запятая по умолчанию зарезервирована под десятичный разделитель. Поэтому в качестве разделителя используется точка с запятой.

CSV файл удобен тем, что для его открытия не нужен дополнительный инструмент (достаточно обычного текстового редактора), преобразовать БДС в csv формат можно, к примеру, через Microsoft Excel (рис. 6).

фрагмент Базы данных симптомов

|   |   |   |    |       |   |    |      |
|---|---|---|----|-------|---|----|------|
| Клинические признаки поражения суставов (припухлость, воспаление) | 1 Крупный сустав (покраснение)          | + | 0  | 0     | + | 10 | 7,69 |
| Клинические признаки поражения суставов (припухлость, воспаление) | 2-10 крупных суставов                   | + | 10 | 5,43  | - | 0  | 0    |
| Клинические признаки поражения суставов (припухлость, воспаление) | 1-3 мелких сустава (без учета крупных)  | + | 20 | 10,87 | - | 0  | 0    |
| Клинические признаки поражения суставов (припухлость, воспаление) | 4-10 мелких сустава (без учета крупных) | + | 30 | 16,3  | - | 0  | 0    |

результатирующий csv файл

```
Клинические признаки поражения суставов (припухлость, воспаление);1 Крупный сустав (покраснение);+;0;0;+;10;7,69;
Клинические признаки поражения суставов (припухлость, воспаление);2-10 крупных суставов;+;10;5,43;-;0;0;
Клинические признаки поражения суставов (припухлость, воспаление);1-3 мелких сустава (без учета крупных);+;20;10,87;-;0;0;
Клинические признаки поражения суставов (припухлость, воспаление);4-10 мелких сустава (без учета крупных);+;30;16,3;-;0;0;
```

Рис. 6 Фрагмент базы данных симптомов и соответствующий ей csv файл.

Экспертная система обращается к файлу и извлекает из него данные по каждому из симптомов построчно. База данных симптомов является *внешней*, заменяемой частью системы, выбор которой осуществляется при запуске программы. Таким образом, разработанная ЭС может решать задачи, связанные с диагностикой широкого круга заболеваний.

## 2.1. Структура программы

Представленная медицинская ЭС была реализована на языке программирования C# в среде разработки MS Visual Studio 2013. Структура программы, состоящей из связанных классов и структур, представлена на рис. 7.

Класс *Calculation* содержит в себе основные методы работы программы: метод *LoadBDS*, предназначенный для выгрузки данных из базы данных симптомов, метод *CreateUserInterface*, включающий в себя адаптируемый пользовательский интерфейс для взаимодействия

пользователя и программы и метод *Calculate*, содержащий механизм диагностирования на основе дифференциально-диагностических признаков.

Класс *Illness* предназначен для хранения информации о диагностируемых заболеваниях, такой как: список характерных симптомов, наличие или отсутствие симптома у данной разновидности заболевания, частота встречаемости, весовые коэффициенты. Эти данные формируются после обработки базы данных симптомов и хранятся в программе в виде списков соответствующих типов. В классе *Illness* также формируются эталонные вектора соответствия заболеваний. Информация о симптомах, на которую опирается данный класс, хранится в структуре *Symptom*, состоящей из двух свойств – *name* и *category*. В строковом свойстве *name* хранится название симптома, а в свойстве *category*, являющемся экземпляром одноименной структуры - ссылка на категорию, к которой принадлежит данный симптом. Структура *Category* является вспомогательной, в ней хранится информация об относящихся к ней симптомах.

Для хранения информации о векторе соответствия больного сформирован класс *Patient*. Помимо списка, содержащего данные об анкетировании пациента (наличие или отсутствие симптома), данный класс включает в себя словарь *Diagnose<Illness, double>*, состоящий из пар <ключ, значение>, где <ключ> - заболевание, <значение> - диагностический индекс, т.е. вероятность наличия заболевания в процентном отношении (см. рис. 7).



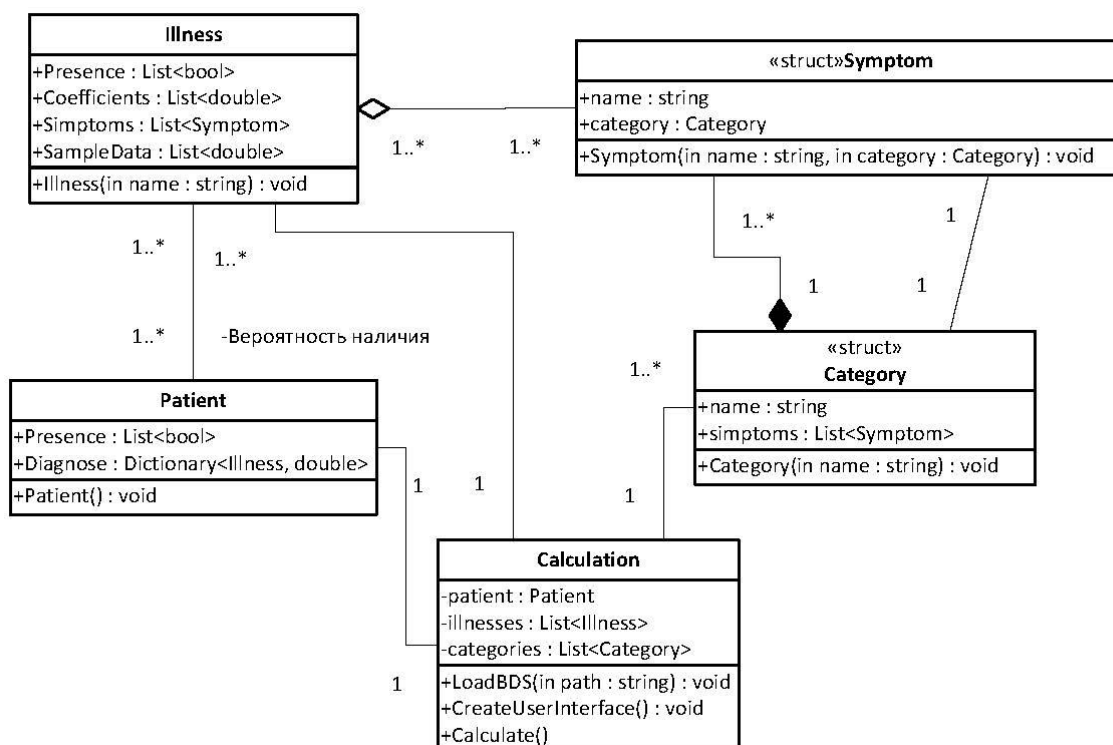


Рис.7. UML-диаграмма структуры программы.

## 2.3. Пользовательский интерфейс

Созданная экспертная система работает на базе совокупности дифференциально-диагностических признаков в форме *вопрос-ответ* [16]. На основе эталонных ответов строятся эталонные матрицы  $A$  логического типа (ячейка принимает значение *true*, если данное заболевание имеет данный симптом, и значение *false* - в противном случае) [15].

Работа программы начинается с выбора через диалоговое окно открытия csv файла, соответствующего базе данных симптомов диагностируемых заболеваний. По перечню содержащихся в нем симптомов формируется интерфейс пользователя (рис. 8), представляющий собой совокупность элементов управления *Checkbox*, разбитых на соответствующие категории. Далее врач отвечает на вопросы о наличии или отсутствии

симптомов у пациента, расставляя «галочки» в форме опросника. Так формируется матрица пациента  $B$  логического типа. Далее матрица пациента сравнивается с каждой из эталонных матриц на наличие совпадений, и для каждого заболевания рассчитывается диагностический индекс  $D$  по формуле (1).

Конечный результат, содержащий возможные заболевания пациента в процентном отношении, выводится на экран и сохраняется в отдельный файл (см. рис. 9).

| Показательные Анализы   | Переносимые процессы и/или болезни | Пол                |
|---|------------------------------------|--------------------|
| Клинические признаки поражения суставов   | Атака артрита                      | Тесты на РФ и АЦЦП |
| <input type="checkbox"/> 1 Крупный сустав (покраснение)   |                                    |                    |
| <input checked="" type="checkbox"/> 2-10 крупных суставов   |                                    |                    |
| <input type="checkbox"/> 1-3 мелких сустава (без учета крупных)   |                                    |                    |
| <input type="checkbox"/> 4-10 мелких сустава (без учета крупных)  |                                    |                    |
| <input type="checkbox"/> >10 суставов (как минимум 1 мелкий сустав)   |                                    |                    |
| <input checked="" type="checkbox"/> Вовлечение первого плюснефалангового сустава                              |                                    |                    |
| <input type="checkbox"/> Энтезопатии (боль в пятке, боль в проекции большеберцового бугра)                    |                                    |                    |
| <input type="checkbox"/> Воспалительная боль в нижнем отделе спины: сакроилеит                                |                                    |                    |
| <input checked="" type="checkbox"/> Воспалительная боль в нижнем отделе спины: спондилит                      |                                    |                    |
| <input checked="" type="checkbox"/> Воспаление связок и сухожилий в месте их прикрепления к седалищному бугру |                                    |                    |

Результат

Рис. 8. фрагмент программы.

Результаты

Пациент: Иванов Ю.А.  
Врач: Сидорова А.К.  
Дата: 02.06.17

|   | № | Форма заболевания    | Диагностический индекс |
|---|---|----------------------|------------------------|
| ▶ | 1 | Реактивный артрит    | 41,63                  |
|   | 2 | Подагрический артрит | 19,23                  |
|   | 3 | Ревматоидный артрит  | 5,43                   |
| * |   |                      |                        |

Сохранить

Рис. 9. Результат программы.

## Выводы

Была разработана и реализована экспертная система для ранней диагностики различных заболеваний, которая включает в себя механизм обращения к базе данных симптомов, алгоритм формирования входных параметров экспертной системы и метод принятия решений на основе математического аппарата вывода данных.

Все поставленные задачи были выполнены. Данная медицинская экспертная система:

- обладает адаптивным пользовательским интерфейсом, позволяющим применять разработанное программное средство для диагностики широкого круга патологий;
- производит раннюю диагностику заболевания по совокупности дифференциально-диагностических признаков;
- обеспечивает специалиста достоверными результирующими данными о наличии диагностируемого заболевания в процентном соотношении.

## Заключение

Результатом данной работы является разработка экспертной системы, которая проводит раннюю диагностику заболеваний, обеспечивает специалиста достоверными результирующими данными и обладает адаптируемым пользовательским интерфейсом, позволяющим использовать ЭС для диагностики широкого круга заболеваний.

Стоит отметить, что достоверность показателя частоты встречаемости симптома может быть повышена, благодаря увеличению числа выборки, то есть при апробации программы в специализированном медицинском центре по диагностике рассматриваемых заболеваний. Или, как в случаях с компанией Socmedica, путем использования медицинских ресурсов – Объединенной Базы Медицинских Знаний (UMKB). Также возможно совершенствование алгоритма принятия решений. Математический аппарат может быть расширен с помощью внедрения логических правил вывода данных, т.е. использования гибридного алгоритма, при этом, не в ущерб времени работы программного продукта.

## Список литературы

1. Овчаренко С.И., Галецкайте Я.К., Волель Б.А., Пушкарев Д.Ф., Лас Е.А. Типы реагирования на хроническое заболевание у пациентов с хронической обструктивной болезнью легких и ревматоидным артритом // Клиницист. 2013. №1. С. 21-29.
2. Медицинские экспертные системы: что это такое? <https://www.kv.by>
3. Журнал «Компьютерные вести On-line». Вып. 40, 2005 г. «Эволюция экспертных систем. История и перспективы». <http://tpl-it.wikispaces.com/MYCIN>
4. Nedovic Lj., Devedzic V., "Expert Systems in Finance - A Cross-Section of the Field", Expert Systems With Applications, Vol.23, No.1, 2002, pp. 49-66.
5. Kanatov M., Atymtayeva L., Yagaliyeva B. Expert systems for Information Security Management and Audit. Implementation phase issues. // SCIS&ISIS. Kitakyushu, Japan: IEEE, 2014. С. 896-900
6. Tsudik, G. and Summers, R. 1990. AudES - an Expert System for Security Auditing. IBM Los Angeles Scientific Center.
7. Экспертные системы для клинической практики. <https://socmedica.com/>
8. Экспертные медицинские системы. <http://www.absoftsite.com/solutions/medicine/expert-systems>
9. Kuzminov O.M., Fetisova V.I., Sinica I.V. Formalization of clinical information for optimization of diagnostic and treatment process in therapeutic practice // Сетевой журнал «Научный результат». Серия «Медицина и фармация». – Т.2, №2, 2016
10. Тонеева Д.В., Гончарова А.Б. Модель представления медицинских данных по Международной классификации функционирования, ограничений жизнедеятельности и здоровья при ведении первичного

- приема детей раннего возраста с нарушениями // Процессы управления и устойчивость . 2016. Т.3, №1. С. 323-326.
11. Липинский Л.В., Кушнарера Т.В., Дябкин Е.В, Попов Е.А. Гибридный эволюционный алгоритм автоматизированного формирования деревьев принятия решения // Вестник СибГАУ . 2014. № № 5(57). С. С. 85–92 .
  12. Kirillov V., Gladyshev A., Demidchik E. Technology of Creation of an Expert System for Diagnosing Thyroid Pathology Based on a Set of Qualitative Signs of Cell Atypia // Microscopy Research and Technique. 2010. №73. С. 1091–1100.
  13. Тонеева Д.В., Гончарова А.Б., Сергеева Е.И. Алгоритм построения экспертной системы диагностики заболеваний на основе дифференциально-диагностических признаков // Технические науки - от теории к практике: сб. ст. по матер. LXIV междунар. науч.-практ. конф. № 11(59). – Новосибирск: СибАК, 2016. – С. 37-43.
  14. Кириллов В.А., Емельянова О.А Диагностика фолликулярного рака и фолликулярной аденомы щитовидной железы с помощью экспертной системы, построенной на базе совокупности качественных дифференциально-диагностических признаков цитогрaмм // Онкологический журнал. - Т.4 №4. С. 87-94.
  15. Лифантова Е.Е., Гончарова А.Б., Машинский Н.С. Создание системы поддержки принятия решения в медицине для диагностики заболеваний желудочно-кишечного тракта // Процессы управления и устойчивость. Том 3(19) №1, СПб, Издательский дом Федоровой Г.В., 2016, С 312-316.
  16. Кириллов В.А., Гладышев А.О., Демидчик Е.П. Экспертная система для диагностики рака щитовидной железы // Онкологический журнал. 2009. №2. С. 5-11.
  17. Тонеева Д.В., Гончарова А.Б. Экспертная система диагностики заболеваний // EUROPEAN RESEARCH: Сборник статей

- VI Международной научно-практической конференции / Под общ.ред. Г.Ю. Гуляева. Пенза: МЦНС «Наука и просвещение», 2016. С. 34–38.
18. Каратеев Д. Е., Олюнин Ю. А., Лучихина Е. Л. Новые классификационные критерии ревматоидного артрита ACR/EULAR 2010 Шаг вперед к ранней диагностике // Научно-практическая ревматология. 2011. №1. С. 10-15.
19. Барскова В.Г. Диагностика подагрического артрита // РМЖ. 2011. №10. С. 614
20. Казакова Т. В, Рашид М. А., Шостак Н. А., Карпова Н. Ю. Реактивный артрит: клиника, диагностика и лечение // Лечебное дело. 2010. №1. С. 11-22.